



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
-----------------	-------------	----------------------	---------------------	------------------

10/729,833

12/05/2003

Todd D. Wakefield

17354.4.5

8543

22913 7590 02/19/2008

WORKMAN NYDEGGER  
60 EAST SOUTH TEMPLE  
1000 EAGLE GATE TOWER  
SALT LAKE CITY, UT 84111

EXAMINER

DAYE, CHELCIE L

ART UNIT

PAPER NUMBER

2161

MAIL DATE

DELIVERY MODE

02/19/2008

PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

Mr

<b>Office Action Summary</b>	<b>Application No.</b> 10/729,833	<b>Applicant(s)</b> WAKEFIELD ET AL.	
	<b>Examiner</b> Chelcie Daye	<b>Art Unit</b> 2161	

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --**

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 26 November 2007.
- 2a) ☒ This action is **FINAL**.                      2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-30 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-30 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \*    c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)            | 4) <input type="checkbox"/> Interview Summary (PTO-413)           |
| 2) <input type="checkbox"/> Notice of Draftperson's Patent Drawing Review (PTO-948)    | Paper No(s)/Mail Date. _____                                      |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date <u>10/24/07</u> .  | 6) <input type="checkbox"/> Other: _____                          |

### **DETAILED ACTION**

1. This action is issued in response to applicant's amendment filed November 26, 2007.
2. Claims 1-30 are presented. Claims 29 and 30 are added and none cancelled.
3. Claims 1-30 are pending.

### ***Information Disclosure Statement***

The information disclosure statement (IDS) submitted on 10/24/07 was filed after the mailing date of the application. The submission is in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.

### ***Claim Objections***

4. Claims 1 and 3 are objected to because of the following informalities:
  - Claim 1 – replace the colon (:) with a semicolon (;) at the end of the first limitation and replace the period (.) with a semicolon (;) after the second limitation of the claim.
  - Claim 3 - the "I" within the first line of the claim for the word "In" should be a lowercase.

Appropriate corrections are required.

***Claim Rejections - 35 USC § 112***

5. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

6. Claim 29 is rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

In particular, newly added claim 29 recites "post-processing filters". However, the applicant's specification does not specifically provide support for post-processing filters, as well as the examiner is unable to determine, with respect to the applicant's own disclosure, as to why the filters are only able to be applied post-process. As such, the newly added claim appears to have new subject matter, which is not taught within the specification. Corrections are required.

7. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

8. Claims 2 and 9 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Newly amended claim 2 recites the limitation "the derived source text" within the limitation of the claim. The examiner is unsure as to what "source text" applicant is referring to, since there is no prior mention of such text. Therefore, there is insufficient antecedent basis for this limitation in the claim.

Claim 9 recites the limitation "said the stored integrated data" within the limitation of the claim. The examiner is unsure as to what "stored integrated data" applicant is referring to, since there is no prior mention of such data. It seems as if applicant previously deleted such step of storing within the relied upon amended independent claim 1. Therefore, there is insufficient antecedent basis for this limitation in the claim.

### ***Claim Rejections - 35 USC § 102***

9. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

**10. Claims 1, 8, 10-15, 22, and 24-28 are rejected under 35 U.S.C. 102(b) as being anticipated by Horvitz (US Patent No. 5,864,848) issued January 26, 1999.**

Regarding Claims 1 and 15, Horvitz discloses at product of a process of integration of mixed format data, said product comprising at least one storage media device including a database of integrated data, said product produced by an integration process comprising the steps of:

accessing a database of data records, at least some of the data records containing both structured and unstructured data, the unstructured data of a particular data record including free text related to the structured data of that data record (Fig.5; col.16, lines 38-45, Horvitz)<sup>1</sup>;

using linguistic characteristics of the free text to interpret said free text to extract relational facts from it (col.17, lines 32-40 and col.18, lines 38-49, Horvitz),

producing a set of construed data reflecting at least one relational fact conveyed in the free text, each construed datum being relatable to the structured data of the same data record (Fig.6B; col.17, lines 17-30, Horvitz);

integrating the construed data with the particular structured data to which the construed data relates (Figs.6B-C; col.17, lines 41-65 and col.19, lines 53-59, Horvitz).

Regarding Claims 8 and 22, Horvitz discloses a product of the process wherein the produced file has a format selected from the group of XML, character separated values, spreadsheet formats and file-based database structures (col.17, lines 17-40, Horvitz).

Regarding Claims 10 and 24, Horvitz discloses a product of the process wherein: the process further includes the step of constructing a library containing extracted attributes; and said product includes the constructed library (col.13, lines 10-26, Horvitz).

---

<sup>1</sup> Examiner Notes: See cols.16-17, lines 64-67 and 1-16, respectively; for further disclosure of the

Regarding Claims 11,14,25, and 28, Horvitz discloses a product of the process wherein said interpreting the free text further includes the step of coalescing like relation types (Figs.6A-B, Horvitz).

Regarding Claims 12,13, and 26-27, Horvitz discloses a product of the process wherein said interpreting the free text further includes the step of coalescing like attributes (Figs. 6A-B, Horvitz).

***Claim Rejections - 35 USC § 103***

11. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

**12. Claims 2-5, 16-19, and 29-30 are rejected under 35 U.S.C. 103(a) as being unpatentable over Horvitz (US Patent No. 5,864,848) issued January 26, 1999, in view of Loatman (US Patent No. 4,914,590) filed May 18, 1988.**

Regarding Claims 2,16, and 29, Horvitz discloses all of the claimed subject matter as stated above. However, Horvitz is silent with respect to identifying roles within the parsed text records and applying caseframes, said application of caseframes

---

structured and unstructured data.

Art Unit: 2161

producing attribute extractions, each of said attribute extractions containing attribute information of the derived source text. On the other hand, Loatman discloses identifying roles within the parsed text records and applying caseframes (col.6, lines 35-51, Loatman), said application of caseframes producing attribute extractions, each of said attribute extractions containing attribute information of the derived source text (col.57, lines 20-34, Loatman). It would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate Loatman's teachings into the Horvitz system. A skilled artisan would have been motivated to combine in order to provide an improved approach to understanding natural language. In particular to understanding linguistic domains, as well as understanding a variety of applications such as electronic mail.

Regarding Claims 3,17, and 30, the combination of Horvitz in view of Loatman, disclose a product of the process wherein said interpreting the free text further includes the step of identifying domains of the set of attribute extractions (col.66, lines 23-66, Loatman).

Regarding Claims 4 and 18, the combination of Horvitz in view of Loatman, disclose a product of the process wherein said interpreting the free text further includes the step of assigning domain roles (col.66, lines 5-20, Loatman).



Regarding Claims 5 and 19, the combination of Horvitz in view of Loatman, disclose a product of the process wherein said interpreting the free text further includes the step of producing relation types (col.61, lines 17-24, Loatman).

**13. Claims 6-7, 9, 20-21, and 23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Horvitz (US Patent No. 5,864,848) filed January 31, 1997, in view of Alpha (US Patent No. 6,980,976) filed August 13, 2001.**

Regarding Claims 6 and 20, Horvitz discloses all of the claimed subject matter as stated above. However, Horvitz is silent with respect to the step of creating a new database containing the structured data element produced in said producing; and said product includes the created database. On the other hand, Alpha discloses the step of creating a new database containing the structured data element produced in said producing; and said product includes the created database (col.5, lines 14-32 & 58-67; col.6, lines 1-12, Alpha). It would have been obvious to one of ordinary skill in the art at the time of the invention to incorporate Alpha's teachings into the Horvitz system. A skilled artisan would have been motivated to combine as a way of better structuring and understanding both free-text and structured text within a document.

Regarding Claims 7 and 21, the combination of Horvitz in view of Alpha, disclose a product of the process wherein: said interpreting the free text further includes the step of producing a file containing the structured data element produced in said producing;

Art Unit: 2161

and said product includes the produced file (col.5, lines 14-32 & 58-67; col.6, lines 1-12, Alpha).

Regarding Claims 9 and 23, the combination of Horvitz in view of Alpha, disclose a product of the process wherein said the stored integrated data includes reference information to the original free text for construed data (col.5, lines 29-32, Alpha).

### **ALTERNATE REJECTION**

#### ***Claim Rejections - 35 USC § 102***

14. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

15. **Claims 1-30 are rejected under 35 U.S.C. 102(e) as being anticipated by Lemus (US Patent Application No. 20020156817).**

Regarding Claims 1 and 15, Lemus teaches a product of a process of integration of mixed format data, said product comprising at least one storage media device including a database of integrated data, said product produced by an integration

Art Unit: 2161

process comprising the steps of: accessing a database of data records, at least some of the data records containing both structured data and unstructured data; the unstructured data of a particular data record including free text related to the structured data of that data record (i.e., Lemus teaches a database from which structured data can be accessed; Lemus also teaches accessing unstructured data e.g., text files and converting/integrating the unstructured data into structured data i.e., for use in a database (i.e., converted into **fields/columns/attributes and tuples/records**). The unstructured data e.g., a text sentence is parsed (i.e., separated into components or **atomized** or character separated values) and attributes are extracted. In the process of attributes extraction/filtration, two methods are used. One filtration method is domain/role independent i.e., based on linguistic features e.g., syntactic structure of the sentence (as in Applicant's own specification. Applicant calls it syntactic caseframe. See paragraph [0009] of the instant specification) and the other method is domain (i.e., role) dependent where the attributes are filtered according to the context/role/domain. The extracted and filtered unstructured data now needs to be put in a database. Therefore the attributes/fields are normalized, identified and extracted attributes/fields are mapped to field/attribute of relation i.e. (EIS i.e., a relational database e.g., Sybase) using database schema information. All of the attributes are **coalesced/categorized** i.e. like attributes are combined together.) (paragraph [0013]; [0030]; [0031 ]; [0038]; [0051 ]; [0052]; [0056]; [0061]; [0063]; [0064]; [0065]; [0066]; [0072]; [0073];[0074];[0083];[0090];[0103];[0104];[0106]); using linguistic characteristics of the free text to interpret said free text to extract relational facts from it; producing a set of construed data reflecting at least one relational fact conveyed in the free text, each construed datum being relatable to the structured data of the same record (See

Art Unit: 2161

explanations above)(paragraph [0013]; [0030]; [0031]; [0038]; [0051]; [0052]; [0056]; [0061]; [0063]; [0064]; [0065]; [0066]; [0072]; [0073]; [0074]; [0083];[0090];[0103];[0104];[0106]); integrating the construed data with the particular structured data to which the construed data relates (See explanations above)(paragraph [0013]; [0030]; [0031 ]; [0038]; [0051 ]; [0052]; [0056]; [0061]; [0063]; [0064]; [0065]; [0066]; [0072]; [0073]; [0074]; [0083]; [0090];[0103];[0104];[0106])).

As to claims 2-14 and 16-30, the limitations of these claims are either addressed or rejected in rejected claim 1 above.

### ***Other Prior Art Made of Record***

**Saldanha et al.** (US Patent No. 6,714,939) discloses a system and method for converting plain text into structured data. Parse trees for the plain text are generated based on the grammar of a natural language, the parse trees are mapped on to instance trees generated based on an application-specific model.

**Friedman et al.** (US Patent No. 6,182,029) discloses a computerized method for extracting information from natural-language text data includes parsing the text data to determine the grammatical structure of the text data and regularizing the parsed text data to form structured word terms. Preferably, the structured output is mapped back to the words in the original sentences of the text data input using XML tags.

**Chen et al.** (US Patent Application No. 2003/0149586) discloses a method for inputting structured data in a first format and unstructured data from a real process from a service or manufacturing operation. The method converts the unstructured information into a second

structured format. In some embodiments, there may not be any unstructured data. The method combines the structured data in first format and structured data in second format.

**Soderland** ("Learning Text Analysis Rules for Domain-Specific Natural Language Processing") teaches a system that automatically induces domain-specific text analysis rules from training examples.

**Soderland** ("Learning Information Extraction Rules for Semi-Structured and Free Text") teaches an information extraction system that is designed to handle text styles ranging from highly structured to free text.

**Volantia** ("About Volantia") describes the service offered by Volantia.

### ***Response to Arguments***

Applicant's arguments with respect to the newly amended claims have been considered but are moot in view of the new ground(s) of rejection.

### ***Conclusion***

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action.

Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the

Art Unit: 2161

shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

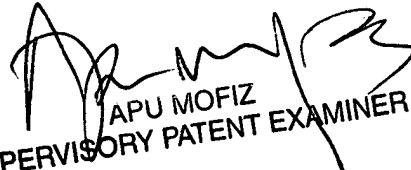
***Points of Contact***

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Chelcie Daye whose telephone number is 571-272-3891. The examiner can normally be reached on M-F, 7:00 - 4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Apu Mofiz can be reached on 571-272-4080. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Chelcie Daye  
Patent Examiner  
Technology Center 2100  
February 12, 2008

  
APU MOFIZ  
SUPERVISORY PATENT EXAMINER